Rotations in three, four, and five dimensions

jason hanson

January 18, 2013

Abstract

The geometry of rotations in dimensions 3, 4, and 5 is discussed using the matrix exponential map. Explicit closed formulas for the exponential of an antisymmetric matrix, as well as the logarithm of a rotation, are given for these dimensions.

1   Introductory remarks

For representing rotations in three dimensions, one usually considers two choices: Euler angles, or quaternions. For aerospace navigation, Euler angles are a natural choice, since the familiar maneuvers of yaw, pitch, and roll are coordinate axis rotations. However, from a global point of view (as opposed to the local viewpoint of a pilot), Euler angles are unnatural and awkward. In particular, the decomposition of a rotation about a noncoordinate axis into Euler angles is not unique and depends on the sequence of axis rotations. On the other hand, quaternions can be used to represent a rotation in a more convenient manner: rotation about an arbitrary axis can be written in a simple form, and given a rotation, the extraction of the rotation axis and angle is a relatively straight–forward process. The price paid for this convenience, of course, is the introduction of a less familiar Mathematical apparatus.

In less common use is a third choice: the representation of a rotation by a linear map, or matrix. Fundamentaly, the Euler angle representation is a matrix representation, but one only uses the representation of coordinate axis rotations, which have a very simple form. In some texts, one may encounter a matrix formula for rotation about an arbitrary axis, but usually this is

1

presented in the form of a rather unappealing and unyieldy 3×3 matrix. In this article, we will recast this matrix in a more appealing form by using some simpler linear maps. As a natural byproduct of this approach, we will obtain a simple way of extracting the rotation axis and angle, but using a more familiar apparatus than quaternions. Moreover, we will be able to extend the constructions to higher dimensions. While one may use quaternions to represent rotations in four dimensions (see [3]), one cannot go beyond.

2 Rotations in three dimensions

A rotation R in three dimensional space is specified by a (real) scalar, the *rotation angle* θ, and a (real) unit vector $\mathbf{u}$, the *rotation axis*. The effect of R on any vector $\mathbf{v}$ is given by *Rodrigues' Rotation Formula:*

$$R\mathbf{v} = \cos\theta\,\mathbf{v} + (1 - \cos\theta)(\mathbf{u} \cdot \mathbf{v})\mathbf{u} + \sin\theta\,\mathbf{u} \times \mathbf{v}. \tag{1}$$

Geometrically, if we write $\mathbf{v}$ in terms of its components parallel and perpendicular to $\mathbf{u}$ (respectively): $\mathbf{v} = \mathbf{v}_\parallel + \mathbf{v}_\perp$, then $R\mathbf{v}_\parallel = \mathbf{v}_\parallel$, and $R\mathbf{v}_\perp$ is the rotation of $\mathbf{v}_\perp$ in the plane orthogonal to $\mathbf{u}$ by θ radians. It is important to note that this is *counterclockwise* rotation about $\mathbf{u}$; indeed, using the standard three–dimensional vector identity $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$, one computes $\mathbf{v}_\perp \times (\mathbf{u} \times \mathbf{v}) = |\mathbf{v}_\perp|^2\mathbf{u}$, which is a positive multiple of $\mathbf{u}$.

2.1 Rotation as a linear map

We may express equation (1) as a 3×3 matrix, or equivalently, a linear map. Let $proj_\mathbf{u}$ denote the linear map $proj_\mathbf{u}(\mathbf{v}) \doteq (\mathbf{u} \cdot \mathbf{v})\mathbf{u}$ (we are still assuming that $|\mathbf{u}| = 1$); that is, $proj_\mathbf{u}$ is orthogonal projection onto $\mathbf{u}$. In addition, let $\Lambda_\mathbf{u}$ denote the linear map $\Lambda_\mathbf{u}(\mathbf{v}) \doteq \mathbf{u} \times \mathbf{v}$. Explicitly as matrices,

$$proj_\mathbf{u} = \begin{pmatrix} u_1^2 & u_1 u_2 & u_1 u_3 \\ u_1 u_2 & u_2^2 & u_2 u_3 \\ u_1 u_3 & u_2 u_3 & u_3^2 \end{pmatrix} \quad \text{and} \quad \Lambda_\mathbf{u} = \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix}$$

where $\mathbf{u} = (u_1, u_2, u_3)$. Using these maps, equation (1) becomes

$$R = \cos\theta\,I + (1 - \cos\theta)\,proj_\mathbf{u} + \sin\theta\,\Lambda_\mathbf{u}. \tag{2}$$

Here I denotes the identity map: $I\mathbf{v} \doteq \mathbf{v}$.

2.2 Finding the angle and axis of a rotation

Conversely, given a rotation R, we may deduce its rotation angle and axis. For the angle, we take the trace of R in equation (2) to get $tr(R) = 1 + 2\cos\theta$. Thus

$$\theta = \cos^{-1} \tfrac{1}{2} \big(tr(R) - 1 \big). \tag{3}$$

For the axis, we make use of the fact that I and $proj_{\mathbf{u}}$ are symmetric, while $\Lambda_{\mathbf{u}}$ is antisymmetric, so that equation (2) also yields

$$\tfrac{1}{2}(R - R^t) = \sin\theta \, \Lambda_{\mathbf{u}}, \tag{4}$$

where R^t denotes the tranpose of R. The components of $\mathbf{u}$ can be read directly off of $\Lambda_{\mathbf{u}} = (R - R^t)/2\sin\theta$ (or more practically, since $2\sin\theta \, \Lambda_{\mathbf{u}} = \Lambda_{2\sin\theta \, \mathbf{u}}$, we can just read off the components of $2\sin\theta \, \mathbf{u}$ from $R - R^t$, and then normalize to get $\mathbf{u}$).

Two remarks are in order. First, the rotation angle θ determined by equation (3) is in the range $0 \le \theta \le \pi$, whereas there is no such restriction in equation (2). If the original rotation angle is greater than π (but less that 2π), then the axis determined from (4) will be opposite of the original rotation axis. Second, we can only use equation (4) to determine the rotation axis if $0 < \theta < \pi$. In the case $\theta = 0$, the rotation is the identity. And in the case $\theta = \pi$, the rotation is reflection in the plane orthogonal to the vector $\mathbf{u}$, which we can deduce up to sign using $proj_{\mathbf{u}} = \tfrac{1}{2}(I + R)$.

2.3 Rotation as an exponential

Recall that the Taylor series expansion of the exponential function is given by $e^x = \sum_{k\ge0} x^k/k!$, which converges for all real numbers x. Similary, the exponential of a square matrix A is *defined* by the formula

$$\exp(A) \doteq \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

Here we set $A^0 \doteq I$. It can be shown that this summation converges for any matrix square A (see for example, [1]).

It is known from the theory of Lie groups that every rotation is the exponential of an antisymmetric matrix. We can verify this directly in three dimensions. Let $\mathbf{u}$ be a three–dimensional unit vector. Starting from the

matrix representation of $\Lambda_{\mathbf{u}}$, or from vector identities, one computes that $\Lambda_{\mathbf{u}}^2 = -(I - proj_{\mathbf{u}})$ and $\Lambda_{\mathbf{u}}^3 = -\Lambda_{\mathbf{u}}$. In particular, $-\Lambda_{\mathbf{u}}^2$ is orthogonal projection onto the plane orthogonal to $\mathbf{u}$. It follows that $\Lambda_{\mathbf{u}}^{2p} = (-1)^p(I - proj_{\mathbf{u}})$ for all $p \geq 1$, and $\Lambda_{\mathbf{u}}^{2p+1} = (-1)^p\Lambda_{\mathbf{u}}$ for all $p \geq 0$. We then compute

$$\exp(\theta\Lambda_{\mathbf{u}}) = \sum_{k=0}^{\infty} \frac{\theta^k}{k!} \Lambda_{\mathbf{u}}^k = I + \sum_{p=1}^{\infty} \frac{\theta^{2p}}{(2p)!} \Lambda_{\mathbf{u}}^{2p} + \sum_{p=0}^{\infty} \frac{\theta^{2p+1}}{(2p+1)!} \Lambda_{\mathbf{u}}^{2p+1}$$

$$= I + \left[\sum_{p=1}^{\infty} \frac{(-1)^p}{(2p)!} \theta^{2p}\right] (I - proj_{\mathbf{u}}) + \left[\sum_{p=0}^{\infty} \frac{(-1)^p}{(2p+1)!} \theta^{2p+1}\right] \Lambda_{\mathbf{u}}$$

$$= I + (\cos\theta - 1)(I - proj_{\mathbf{u}}) + \sin\theta\, \Lambda_{\mathbf{u}}$$

by virtue of the Taylor series expansions for sine and cosine. This last expression is just a rearrangement of equation (2). That is, $\exp(\theta\Lambda_{\mathbf{u}})$ is counterclockwise rotation by θ radians about $\mathbf{u}$.

3 Simple rotations

We now extend the above constructions used in three dimensions to any dimension. To do this, we define two special linear maps. If $\mathbf{u}$ and $\mathbf{v}$ are two vectors in n–dimensional space, their **outer product** is the linear map $\mathbf{u} \otimes \mathbf{v}$ such that

$$(\mathbf{u} \otimes \mathbf{v})(\mathbf{w}) = (\mathbf{v} \cdot \mathbf{w})\mathbf{v}$$

for all vectors $\mathbf{w}$. This is a generalization of the projection operation: if $|\mathbf{u}| = 1$, then $proj_{\mathbf{u}} = \mathbf{u} \otimes \mathbf{u}$. The **wedge product** of $\mathbf{u}$ and $\mathbf{v}$ is defined as

$$\mathbf{u} \wedge \mathbf{v} \doteq \mathbf{u} \otimes \mathbf{v} - \mathbf{v} \otimes \mathbf{u}.$$

The wedge product will serve as our generalization of $\Lambda_{\mathbf{u}}$; in fact in three dimensions, we have $\mathbf{u} \wedge \mathbf{v} = -\Lambda_{\mathbf{u}\times\mathbf{v}}$.

3.1 Fundamental computations

The following properties of the outer and wedge products are readily verified from the definitions.

Lemma 3.1. *For all vectors* $\mathbf{a}$, $\mathbf{b}$, $\mathbf{u}$, $\mathbf{v}$, *we have*

i. $tr(\mathbf{u} \otimes \mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$

ii. $(\mathbf{a} \otimes \mathbf{b})(\mathbf{u} \otimes \mathbf{v}) = (\mathbf{b} \cdot \mathbf{u})(\mathbf{a} \otimes \mathbf{v}).$

iii. $|\mathbf{u} \wedge \mathbf{v}|^2 proj_{\mathbf{uv}} = |\mathbf{v}|^2 \mathbf{u} \otimes \mathbf{u} - (\mathbf{u} \cdot \mathbf{v})(\mathbf{u} \otimes \mathbf{v} + \mathbf{v} \otimes \mathbf{u}) + |\mathbf{u}|^2 \mathbf{v} \otimes \mathbf{v}.$

iv. $(\mathbf{u} \wedge \mathbf{v})^2 = -|\mathbf{u} \wedge \mathbf{v}|^2 proj_{\mathbf{uv}}$

v. $(\mathbf{u} \wedge \mathbf{v})^3 = -|\mathbf{u} \wedge \mathbf{v}|^2 \mathbf{u} \wedge \mathbf{v}$

where $|\mathbf{u} \wedge \mathbf{v}| \doteq \sqrt{|\mathbf{u}|^2 |\mathbf{v}|^2 - (\mathbf{u} \cdot \mathbf{v})^2}$, *and* $proj_{\mathbf{uv}}$ *is orthogonal projection onto the plane spanned by* $\mathbf{u}$ *and* $\mathbf{v}$. $\square$

We remark that, as the notation indicates, $|\mathbf{u} \wedge \mathbf{v}|$ is the matrix norm of $\mathbf{u} \wedge \mathbf{v}$. For antisymmetric (or symmetric) matrices A and B, one usually defines their inner product as $A \cdot B \doteq \frac{1}{2} tr(AB^t)$. In this sense, $|\mathbf{u} \wedge \mathbf{v}|^2 = (\mathbf{u} \wedge \mathbf{v}) \cdot (\mathbf{u} \wedge \mathbf{v})$. In particular, $\mathbf{u} \wedge \mathbf{v} = 0$ if and only if $|\mathbf{u} \wedge \mathbf{v}| = 0$. Note that we may also write $|\mathbf{u} \wedge \mathbf{v}| = |\mathbf{u}| \, |\mathbf{v}| \sin \phi$, where ϕ is the angle between $\mathbf{u}$ and $\mathbf{v}$.

Using lemma 3.1(iv) and (v), we may again sum the Taylor series as we did in section 2.3 to compute the exponential of $\mathbf{u} \wedge \mathbf{v}$.

Theorem 3.2. *For all vectors* $\mathbf{u}$, $\mathbf{v}$ *with* $|\mathbf{u} \wedge \mathbf{v}| = 1$, *and any real scalar* θ,

$$\exp(\theta \, \mathbf{u} \wedge \mathbf{v}) = I - (1 - \cos \theta) \, proj_{\mathbf{uv}} + \sin \theta \, \mathbf{u} \wedge \mathbf{v}. \quad \square$$

3.2   Rotation in a two–plane

In any dimension greater than one, a (linear) *two–plane* is spanned by two linearly independent vectors. We will call a linear map corresponding to rotation in a two–plane, and that fixes the orthogonal complement, a **simple rotation**.

Lemma 3.3. *Given vectors* $\mathbf{u}$, $\mathbf{v}$, *if* $\mathbf{a} = \alpha \mathbf{u} + \beta \mathbf{v}$ *and* $\mathbf{b} = \gamma \mathbf{u} + \delta \mathbf{v}$ *for scalars* α, β, γ, δ, *then* $\mathbf{a} \wedge \mathbf{b} = (\alpha \delta - \beta \gamma) \, \mathbf{u} \wedge \mathbf{v}$.

Proof. Use the linearity and antisymmetry of the wedge product. $\square$

Lemma 3.4. *The vectors* $\mathbf{u}$, $\mathbf{v}$ *span a two–plane if and only if* $\mathbf{u} \wedge \mathbf{v} \neq 0$. *Moreover, the vectors* $\mathbf{a}$, $\mathbf{b}$ *span the same two–plane if and only if* $\mathbf{a} \wedge \mathbf{b}$ *is a nonzero multiple of* $\mathbf{u} \wedge \mathbf{v}$.

Proof. The first statement follows from the remarks after lemma 3.1. The necessity of the second follows from the previous lemma; for the sufficiency, use lemma 3.1(iv) to show that $proj_{\mathbf{ab}} = proj_{\mathbf{uv}}$. $\qquad\square$

An *orientation* for the two–plane spanned by $\mathbf{u}$, $\mathbf{v}$ is a choice of ordering of the vectors, say $(\mathbf{u}, \mathbf{v})$. By lemma 3.3, a pair of vectors $(\mathbf{a}, \mathbf{b})$ from this two–plane has the same orientation if $\mathbf{a} \wedge \mathbf{b}$ is a positive multiple of $\mathbf{u} \wedge \mathbf{v}$.

Theorem 3.5. *If $|\mathbf{u} \wedge \mathbf{v}| = 1$, then $\exp(\theta\,\mathbf{u} \wedge \mathbf{v})$ is simple counterclockwise rotation by θ radians in the oriented two–plane spanned by $(\mathbf{u}, \mathbf{v})$.*

Proof. By lemma 3.4, we may assume $\mathbf{u}$ and $\mathbf{v}$ are mutually orthonormal. Using theorem 3.2, we then compute $\exp(\theta\,\mathbf{u} \wedge \mathbf{v})\mathbf{u} = \cos\theta\,\mathbf{u} - \sin\theta\,\mathbf{v}$, and $\exp(\theta\,\mathbf{u} \wedge \mathbf{v})\mathbf{v} = \sin\theta\,\mathbf{u} + \cos\theta\,\mathbf{v}$. Moreover, if $\mathbf{w}$ is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$, then $\exp(\theta\,\mathbf{u} \wedge \mathbf{v})\mathbf{w} = \mathbf{0}$. $\qquad\square$

3.3 Logarithm of a simple rotation

From theorem 3.5, we see that every simple rotation R can be expressed as the exponential of the wedge product of two vectors, say $R = \exp(\mathbf{u} \wedge \mathbf{v})$. In this case, we refer to $\mathbf{u} \wedge \mathbf{w}$ as a *logarithm* of R. Logarithms of rotations are not unique, as the sine and cosine functions are periodic. However, for simple rotations by angles less that π, there is a canonical choice for logarithm.

Theorem 3.6. *If R is a simple rotation in dimension n, then the angle of rotation is given by $\theta = \cos^{-1}\frac{1}{2}\big(tr(R) - n + 2\big)$. Moreover, if $0 < \theta < \pi$, then $R = \exp(f)$, where $f = (\theta/2\sin\theta)(R - R^t)$.*

Proof. From lemma 3.1, $tr(proj_{\mathbf{uv}}) = 2$ and $tr(\mathbf{u} \wedge \mathbf{v}) = 0$. Thus from theorem 3.2, $tr\big(\exp(\theta\,\mathbf{u} \wedge \mathbf{v})\big) = n - 2(1 - \cos\theta)$. $\qquad\square$

If $0 < \theta < \pi$, we write $\log(R) = f$. If $\theta = 0$, we simply take $\log(R) = I$. And if $\theta = \pi$, then we are only able to deduce the projection onto the two–plane: $proj_{\mathbf{uv}} = \frac{1}{2}(I - R)$; R is multiplication by -1 in this two–plane.

4 Decomposing an antisymmetric matrix

In two and three dimensions, every rotation is simple. In higher dimensions, one must compose simple rotations to obtain a general rotation. If the two–planes of the constituent simple rotations are mutually orthogonal (from Lie

Theory, this can always be done), we may obtain an explicit formula for the composite rotation. We will do this in the next section, although we outline the general procedure here.

Lemma 4.1. *The two–planes spanned by the pairs* $\mathbf{u}$, $\mathbf{v}$ *and* $\mathbf{a}$, $\mathbf{b}$ *are orthogonal if and only if* $(\mathbf{u} \wedge \mathbf{v})(\mathbf{a} \wedge \mathbf{b}) = 0$.

Proof. Since $\mathbf{u} \cdot \mathbf{a} = \mathbf{u} \cdot \mathbf{b} = \mathbf{v} \cdot \mathbf{a} = \mathbf{v} \cdot \mathbf{b} = 0$, lemma 3.1(ii) implies that $(\mathbf{u} \wedge \mathbf{v})(\mathbf{a} \wedge \mathbf{b}) = 0$. Conversely, if $(\mathbf{u} \wedge \mathbf{v})(\mathbf{a} \wedge \mathbf{b}) = 0$, then item lemma 3.1(iv) implies that $proj_{\mathbf{uv}} proj_{\mathbf{ab}} = 0$, and the two–planes are orthogonal. ☐

Suppose that f_j, $1 \leq j \leq m$, are wedge products with mutually orthgonal two–planes, lemma 4.1 then implies that $(\sum_j f_j)^k = \sum_j f_j^k$ for all $k > 0$. By summing the Taylor series, we thus find $\exp(\sum_j f_j) = I + \sum_j (\exp(f_j) - I)$. Theorem 3.2 can be then used to obtain an explicit formula for the latter.

If we are able to write an antisymmetric matix f as a the sum of wedge products f_j whose two–planes are mutually orthogonal, we will call the decomposition $f = \sum_j f_j$ an **orthogonal decomposition** of f. By dimensional considerations, there can be at most $\lfloor n/2 \rfloor$ (where $\lfloor x \rfloor$ denotes the integer part of x) nozero summands in the orthogonal decomposition of a $n \times n$ antisymmetric matrix.

4.1   Existence of orthogonal decomposition

Up to this point, we have used only elementary considerations; however the the following lemma makes use of the *spectral theorem for real $n \times n$ antisymmetric matrices* A: there exists an orthogonal matrix P such that $A = PBP^t$, where B has 2×2 blocks of the form $\left(\begin{smallmatrix} 0 & \lambda \\ -\lambda & 0 \end{smallmatrix} \right)$ along the diagonal; if n is odd, B has at least one additional row and column of zeros. See [2].

Theorem 4.2. *Every real antisymmetric matrix, is the sum of wedge products whose two–planes are mutually orthogonal.*

Proof. Let B and Q be as above; and let $\mathbf{e}_j$ denote the n–dimensional vector whose j–th component is unity, and all other components are zero. We may write $B = \sum_j \lambda_j \mathbf{e}_{j-1} \wedge \mathbf{e}_{j+1}$. Thus B is evidently the sum of wedge products with mutually orthogonal two–planes; and since Q is orthogonal, $A = QBQ^T = \sum_j \lambda_j (Q\mathbf{e}_{j-1}) \wedge (Q\mathbf{e}_{j+1})$ is also. ☐

Corollary 4.3. *Every real antisymmetric $n \times n$ matrix, with $n = 4, 5$, is the sum of at most two wedge products whose two–planes are orthogonal.* ☐

4.2 Finding an orthogonal decomposition

Theorem 4.4. *Let f be a real $n \times n$ antisymmetric matrix, with $n = 4, 5$, then*

$$\Delta \doteq tr(f^4) - \tfrac{1}{4}tr^2(f^2) \quad and \quad \theta_\pm^2 \doteq -\tfrac{1}{4}tr(f^2) \pm \tfrac{1}{2}\sqrt{\Delta}$$

are nonnegative. Moreover if $\Delta \neq 0$, then we have the orthogonal decomposition $f = f_+ + f_-$ where

$$f_\pm = \mp \frac{1}{\sqrt{\Delta}}(\theta_\mp^2 f + f^3)$$

are wedge products with $f_+ f_- = 0$ and $|f_\pm| = \theta_\pm \doteq \sqrt{\theta_\pm^2}$. This decomposition is unique up to reordering of summands.

Proof. By corollary 4.3, we may write f as the sum of wedge products $f = f_+ + f_-$, with $f_+ f_- = 0$. However from lemma 3.1, $f_\pm^2 = -|f_\pm|^2 proj_\pm$, where $proj_\pm$ denotes orthogonal projection onto the two–plane of the wedge product $f_\pm$. Therefore, we have $f^2 = -|f_+|^2 proj_+ - |f_-|^2 proj_-$; and by taking the trace, we get the equation $(*)$ $|f_+|^2 + |f_-|^2 = -\tfrac{1}{2} tr(f^2)$. Similarly, as the square of a projection is itself, one computes the trace of f^4 to obtain the equation: $(**)$ $|f_+|^4 + |f_-|^4 = \tfrac{1}{2} tr(f^4)$. Setting $x \doteq |f_+|^2$ and $y \doteq |f_-|^2$, these two equations give us a system of two algebraic equations in the variables x, y. Using $(*)$ to eliminate y in $(**)$, we arive at the quadratic equation $x^2 + \tfrac{1}{2}tr(f^2)x + \tfrac{1}{8}tr^2(f^2) - \tfrac{1}{4}tr(f^4) = 0$ with discriminant Δ, and whose solutions are $x = \theta_\pm^2$. With possible reindexing, we may take $x = \theta_+^2$ and $y = \theta_-^2$. Note that $(*)$ and $(**)$ also imply that $\Delta = 2(|f_+|^4 + |f_-|^4) - (|f_+|^2 + |f_-|^2)^2 = (|f_+|^2 - |f_-|^2)^2 \geq 0$. In addition, $\theta_\pm^2 = |f_\pm|^2 \geq 0$.

To obtain the formulae for $f_\pm$, we use lemma 3.1(v) to get $f^3 = f_+^3 + f_-^3 = -\theta_+^2 f_+ - \theta_-^2 f_-$. This leads to the Vandermonde system

$$\begin{pmatrix} 1 & 1 \\ \theta_+^2 & \theta_-^2 \end{pmatrix} \begin{pmatrix} f_+ \\ f_- \end{pmatrix} = \begin{pmatrix} f \\ -f^3 \end{pmatrix}.$$

The determinant of the 2×2 matrix on the left is $\theta_-^2 - \theta_+^2 = -\sqrt{\Delta}$. Since we are assuming $\Delta > 0$, we may invert the system to get the stated (necessarily unique) solutions for f_+, f_-. $\square$

In the case when $\Delta = 0$ (so that $\theta_+^2 = \theta_-^2$), an orthogonal decomposition exists, but will not be unique. This stems from the fact that an eigenspace of dimension greater than unity has no natural choice of basis.

5 Rotations in four and five dimensions

As we mentioned previously, every rotation is necessarily the exponential of an antisymmetric matrix; see [1]. We use this fact, coupled with orthogonal decomposition, to write down a parametrization for all rotations in four and five dimensions, and to find logarithms for such rotations.

5.1 Rotations

Theorem 5.1. *Let f, Δ, $\theta_\pm$ be as in theorem 4.4. If $\Delta \neq 0$ and $\theta_\pm \neq 0$, then*

$$\exp(f) = I + \frac{1}{\sqrt{\Delta}}(Af + Bf^2 + Cf^3 + Df^4),$$

where

$$A \doteq \frac{\theta_+^2 \sin\theta_-}{\theta_-} - \frac{\theta_-^2 \sin\theta_+}{\theta_+}, \qquad B \doteq \frac{\theta_+^2(1-\cos\theta_-)}{\theta_-^2} - \frac{\theta_-^2(1-\cos\theta_+)}{\theta_+^2},$$

$$C \doteq \frac{\sin\theta_-}{\theta_-} - \frac{\sin\theta_+}{\theta_+}, \qquad D \doteq \frac{1-\cos\theta_-}{\theta_-^2} - \frac{1-\cos\theta_+}{\theta_+^2}.$$

Moreover if $\Delta = 0$ and $\theta_+ \neq 0$, then $\exp(f) = I + \dfrac{\sin\theta_+}{\theta_+}f + \dfrac{1-\cos\theta_+}{\theta_+^2}f^2$.

Proof. By theorem 4.4 we have $f_+ f_- = 0$, so that $(\theta_-^2 f + f^3)(\theta_+^2 f + f^3) = 0$; from which it follows that $f^6 = -\theta_+^2 \theta_-^2 f^2 - (\theta_+^2 + \theta_-^2)f^4$. If $\Delta \neq 0$ and $\theta_+ \neq 0$, then lemma 3.1(iv) implies $proj_+ = -(1/\theta_+^2\Delta)(\theta_+^2 f + f^3)^2 = (1/\theta_+^2\Delta)(\theta_-^2 f^2 + f^4)$, and similarly for $proj_-$. The first stated formula for $\exp(f)$ then follows from theorem 3.2 and the discussion at the beginning of section 4. On the other hand, if $\Delta = 0$ and $\theta_+ \neq 0$, then $\theta_+ = \theta_-$; so that $|f_+| = |f_-|$. Thus $f^3 = f_+^3 + f_-^3 = -|f_+|^2(f_+ + f_-)$ by lemma 3.1(v); whence $f^3 = -\theta_+^2 f$. We may then sum the Taylor series for $\exp(f)$ to obtain the second stated formula. □

The case $\Delta \neq 0$ and $\theta_+ = 0$ cannot occur, since $-tr(f^2) = |f|^2 \geq 0$; if $\theta_- = 0$, then f is already a wedge product, and theorem 3.2 applies. There are no other nontrivial cases. We remark that in the case when $n = 4$, $\Delta \neq 0$, and $\theta_\pm \neq 0$, we can use the identity $proj_+ + proj_- = I$ to eliminate f^4 in the formula for $\exp(f)$.

5.2 Logarithm of a rotation

Theorem 5.2. *Let R be a $n \times n$ rotation matrix, where $n = 4, 5$. Define*

$$\delta \doteq \tfrac{1}{2} tr(R^2) - \tfrac{1}{4} tr^2(R) + \tfrac{1}{2}(n-4) tr(R) - \tfrac{1}{4} n(n-6)$$

$$y_{\pm} \doteq \tfrac{1}{4}\big(tr(R) - n + 4\big) \pm \tfrac{1}{2}\sqrt{\delta}.$$

Then $\delta \geq 0$ and $-1 \leq y_{\pm} \leq 1$. If $\delta \neq 0$ and $y_{\pm} \neq \pm 1$, then $\exp(f) = R$, where f is antisymmetric and has the orthogonal decomposition $f = f_+ + f_-$ with

$$f_{\pm} \doteq \mp \frac{\theta_{\pm}}{2 \sin \theta_{\pm} \sqrt{\delta}}\big(y_{\mp}(R - R^t) - \tfrac{1}{2}(R^2 - R^{2t})\big),$$

where $\theta_{\pm} \doteq \cos^{-1} y_{\pm}$. If $\delta = 0$ and $y_+ \neq \pm 1$, then $f = \dfrac{\theta_+}{2 \sin \theta_+}(R - R^t)$.

Proof. We know that $R = \exp(\theta_+ w_+ + \theta_- w_-)$, where $w_{\pm}$ are wedge products of unit length and whose two–planes are orthogonal, and where $0 \leq \theta_{\pm} \leq \pi$. By theorem 3.2 and the discussion at the beginning of section 4,

$$R = I - (1 - \cos\theta_+) proj_+ + \sin\theta_+ w_+ - (1 - \cos\theta_-) proj_- + \sin\theta_- w_-, \quad (5)$$

where $proj_{\pm} = -w_{\pm}^2$. Using lemma 3.1(iv) and (v), orthogonality, and the double angle formulas $\sin 2\theta = 2 \sin\theta \cos\theta$ and $\cos 2\theta = \cos^2\theta - \sin^2\theta$, one computes

$$\begin{aligned} R^2 = I - (1 - \cos 2\theta_+) proj_+ + \sin 2\theta_+ w_+ \\ - (1 - \cos 2\theta_-) proj_- + \sin 2\theta_- w_-. \end{aligned} \quad (6)$$

Set $y \doteq \cos\theta_+$ and $z \doteq \cos\theta_-$. Since $\cos 2\theta = 2\cos^2\theta - 1$, we get the equations $(*)$ $tr(R) = 2y + 2z + n - 4$ and $(**)$ $tr(R^2) = 4y^2 + 4z^2 + n - 8$; and from these we get $y^2 - \tfrac{1}{2}[tr(R) - n + 4]y + \tfrac{1}{8}[tr^2(R) - tr(R^2) - 2(n-4)tr(R) + n^2 - 7n + 8] = 0$, whose discriminent is δ, and solutions are $y = y_{\pm}$. Moreover, one computes using $(*)$ and $(**)$ that $\delta = (y_+ - y_-)^2$ and we may take $y_{\pm} = \cos\theta_{\pm}$. Thus $\delta \geq 0$ and $-1 \leq y_{\pm} \leq 1$.

Let $f_+ \doteq \theta_+ w_+$ and $f_- \doteq \theta_- w_-$. First assume that $\delta \neq 0$ and $0 < \theta_{\pm} < \pi$ (that is, $y_{\pm} \neq \pm 1$). Equation (5) then implies that $\tfrac{1}{2}(R - R^t) = \sin\theta_+ f_+/\theta_+ + \sin\theta_- f_-/\theta_-$, and equation (6) implies that $\tfrac{1}{2}(R^2 - R^{2t}) = \sin 2\theta_+ f_+/\theta_+ + \sin 2\theta_- f_-/\theta_-$. Using the double angle formula for sine, we may recast these equations as the matrix equation

$$\begin{pmatrix} 1 & 1 \\ 2y_+ & 2y_- \end{pmatrix} \begin{pmatrix} \sin\theta_+ f_+/\theta_+ \\ \sin\theta_- f_-/\theta_- \end{pmatrix} = \begin{pmatrix} (R - R^t)/2 \\ (R^2 - R^{2t})/2 \end{pmatrix}.$$

The determinant of the 2×2 matrix on the left is $-2\sqrt{\delta} \neq 0$, so we can invert to get the first solution stated. Second, if $\delta = 0$, then $y_+ = y_-$. Therefore, equation (5) becomes $R = I - (1 - y_+)(proj_+ + proj_-) + \sin\theta_+ f/\theta_+$, which implies the second solution, provided that $0 < \theta_+ < \pi$. $\square$

In the case when $\delta \neq 0$ and $y_+ = 1$ or $y_- = 1$, R is a simple rotation, and is handled by theorem 3.6. Note that the case $\delta \neq 0$ and $y_+ = -1$ is not possible (otherwise $y_- < -1$). Moreover, from equation (5), we see that if $y_\pm = -1$ (so that $\delta = 0$), then R is multiplication by -1 in the four–plane with orthogonal projection map $(I - R)/2$ — although this is only meaningful in the case $n = 5$. The only other nontrivial case occurs when $\delta \neq 0$ and $y_- = -1$; in which case $R = \exp(f_+) - 2proj_-$, where $f_+ = (\theta_+/2\sin\theta_+)(R - R^t)$ and $proj_- = (\exp(f_+) - R)/2$.

6   Concluding remarks

The formulas given for rotations and their logarithms in three dimensions seem to be well–known, although surprisingly obscure — most likely due to the predominant use of Euler angles. On the other hand, I have not seen the formulas given here for four and five dimensions previously in print, though it strikes me as unlikey that they are original. Quaternionic formulas for the two–planes and rotation angles for a four–dimensional rotation, however, can be found in [3].

The constructions given for the orthogonal decomposition, exponential map, and logarithm map may be extended to higher dimensions. However for the two–plane rotation angles in dimension n, one must solve a polynomial of degree $\lfloor n/2 \rfloor$, which cannot be done analytically for $n \geq 10$. Moreover to find the orthogonal decomposition, one needs to invert a $\lfloor n/2 \rfloor \times \lfloor n/2 \rfloor$ matrix.

References

[1] Morton L. Curtis, *Matrix Groups*, second edition, Springer, 1984.

[2] Felix R. Gantmacher, *Matrix Theory*, Vol. 2, Am. Math. Soc., 2000.

[3] Joel L. Weiner and George R. Wilkens, *Quaternions and rotations in* $\mathbb{E}^4$, Am. Math. Mon. 112, No. 1, 69–76 (2005).